

ORIGINAL RESEARCH

Open access

Foundation Models in Materials Science: Emerging Architectures and Training Paradigms

Hiroshi Tanaka¹, Yuki Sato^{1*}, Kenji Mori², Rina Okabe¹

Abstract

The convergence of large-scale machine learning with materials engineering is reshaping how new materials are conceived, predicted, and realized. Foundation models—pre-trained architectures that learn generalizable, multimodal representations from expansive datasets—are emerging as the computational backbone for next-generation discovery pipelines. This narrative review synthesizes the computational and data-driven ecosystems that have enabled their rise, drawing on advances in materials informatics, graph-based representation learning, and autonomous experimentation. We trace the progression from early machine learning applications in property prediction to scalable graph neural networks that capture atomic-scale interactions with unprecedented fidelity. High-throughput computation and multimodal data integration have created the knowledge bases necessary for training models that generalize across chemical spaces. Central to this evolution are closed-loop systems, where foundation-like models orchestrate active learning, uncertainty-aware selection, and seamless simulation–experiment feedback. Through an original integrative analysis, we identify recurring architectural principles—such as hierarchical graph convolutions, contrastive pre-training, and multi-task optimization—and training paradigms that balance exploration with exploitation in vast design spaces. These elements collectively address longstanding bottlenecks in inverse design, property optimization, and length-scale bridging. Positioned at the interface of computational infrastructure and autonomous discovery, this review provides a systems-level perspective on how foundation models are poised to compress the materials innovation timeline from decades to months, while maintaining rigorous physical grounding.

Keywords Materials informatics, Graph neural networks, Representation learning, Closed-loop discovery, Foundation models, Autonomous laboratories

*Correspondence:

Yuki Sato

yuki.sato@gmail.com

¹ Department of Computational Materials Engineering, Faculty of Engineering, University of Tokyo, Tokyo, Japan

² Department of Data-Driven Materials Discovery, Faculty of Engineering, Kyoto University, Kyoto, Japan

Introduction

Materials science stands at a pivotal inflection point. For decades, the discovery of new materials relied on a combination of serendipity, empirical intuition, and computationally intensive first-principles calculations. The exponential growth of available data—spanning quantum mechanical simulations, experimental measurements, and scientific literature—has rendered traditional workflows insufficient for navigating the combinatorial explosion of chemical and structural possibilities. Machine learning has emerged as the enabling technology to harness this data deluge, shifting the paradigm from hypothesis-driven to data-driven exploration [1, 2].

Early applications of machine learning in materials focused on supervised property prediction using hand-crafted descriptors [1, 3]. These models demonstrated that statistical patterns in compositional and structural features could yield

accurate forecasts of formation energies, band gaps, and elastic moduli, often surpassing the speed of density functional theory (DFT) while maintaining comparable accuracy [4, 5]. However, their reliance on domain-specific feature engineering limited transferability across material classes and length scales. The subsequent development of representation learning addressed this limitation by allowing models to automatically discover hierarchical embeddings directly from raw atomic configurations [6, 7].

Graph neural networks (GNNs) represented a breakthrough in this regard. By treating atoms as nodes and interatomic bonds as edges, GNNs naturally encode the relational structure of materials, from molecules to periodic crystals [8, 9]. Foundational architectures such as crystal graph convolutional neural networks demonstrated that message-passing mechanisms could propagate local chemical environments into global property predictions with remarkable precision [9]. Subsequent refinements introduced angular information, long-range interactions, and equivariant features, further enhancing physical fidelity [10, 11]. These models not only improved predictive performance but also revealed emergent capabilities, such as out-of-distribution generalization when trained at scale [12].

Parallel to advances in representation learning, the materials community invested heavily in infrastructure for data generation and curation. High-throughput computational platforms, exemplified by the Materials Project, generated millions of DFT-relaxed structures and associated properties, creating standardized benchmarks for model development [6, 13]. These databases, combined with experimental repositories, enabled the construction of multimodal datasets that capture both simulated and measured observables [4]. The resulting data ecosystems provided the raw material for training larger, more expressive models.

A critical insight from recent literature is that model performance scales predictably with data volume and architectural capacity [12]. Large-scale GNNs trained on hundreds of thousands of structures have achieved formation energy accuracies approaching 10–20 meV/atom, sufficient to guide experimental synthesis with high confidence [12]. This scaling behavior mirrors observations in other domains where foundation models have excelled, suggesting that materials science is on the cusp of analogous capabilities. Foundation models in this context are defined not merely by parameter count but by their ability to serve as versatile backbones: pre-trained on diverse tasks (property prediction, structure generation, dynamics simulation) and fine-tuned for downstream applications with minimal additional data [14].

Training paradigms have evolved in tandem. Self-supervised learning, particularly contrastive and masked reconstruction objectives, has proven effective for learning robust representations without exhaustive labeling [15, 16]. Multi-task learning further enhances generalization by forcing models to jointly predict multiple properties, implicitly capturing underlying physical correlations [17]. Active learning strategies, integrated with uncertainty quantification, close the loop between model predictions and new data acquisition, ensuring efficient exploration of chemical space [18].

The integration of these elements into autonomous systems marks the transition from static predictors to dynamic discovery engines. Self-driving laboratories combine robotic hardware, orchestration software, and intelligent agents to execute closed-loop workflows [19–22]. In these platforms, foundation models act as the cognitive layer: proposing candidates via inverse design [23], selecting experiments via Bayesian optimization or reinforcement learning [24], and interpreting outcomes to refine internal representations. Recent demonstrations have achieved order-of-magnitude accelerations in materials optimization, from thin-film deposition to solid-state synthesis [21, 22]. The emergence of foundation models is best understood as the convergence of architectural, data, and training paradigm innovations (Figure 1).

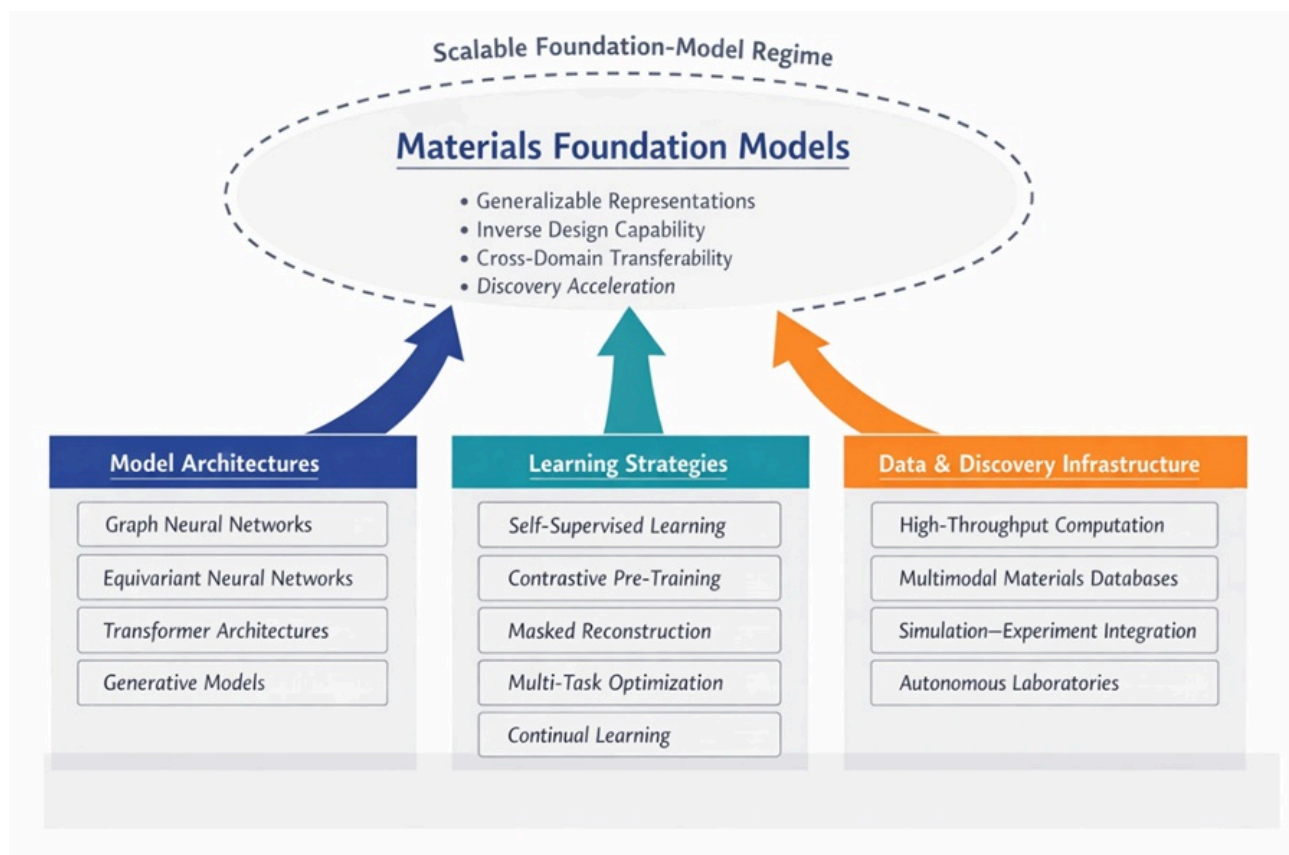


Figure 1. Architectural and Training Paradigm Landscape of Foundation Models in Materials Science

Conceptual landscape mapping the technological convergence enabling foundation models in materials science. The diagram integrates architectural backbones (graph neural networks, equivariant networks), training paradigms (self-supervised learning, multi-task learning, active learning), and infrastructure enablers (high-throughput computation, multimodal datasets, autonomous laboratories). Intersections illustrate how these elements collectively produce scalable, generalizable discovery systems.

This review is structured to provide a comprehensive, infrastructure-oriented synthesis of the field. We first map the broader landscape of computational and data-driven materials engineering, emphasizing the data, models, and workflows that form the substrate for foundation models. We then focus on autonomous and closed-loop systems, where these models are deployed in real-time discovery pipelines. Throughout, we employ original cross-study analysis to highlight integrative themes—such as the interplay between scale and generalizability, the role of multimodal fusion in bridging simulation and experiment, and the necessity of uncertainty-aware decision-making—rather than recapitulating existing taxonomies. By adopting this systems perspective, we aim to equip researchers with a coherent framework for designing, training, and deploying the next generation of foundation models in materials science.

Landscape of Computational & Data-Driven Materials Engineering

Materials informatics and the data-driven ecosystem

The foundation of modern computational materials engineering rests on the systematic collection, curation, and utilization of large-scale datasets. Materials informatics emerged as the discipline that applies data science principles to accelerate discovery, transforming disparate computational and experimental outputs into actionable knowledge [1, 6, 13]. Early

efforts focused on descriptor-based machine learning, where compositional, structural, and electronic features served as inputs to regression or classification models [1, 3]. These approaches demonstrated that even simple statistical models could capture complex structure–property relationships when trained on sufficiently diverse data [2, 7].

The establishment of open-access repositories marked a turning point. Platforms aggregating DFT calculations, experimental phase diagrams, and spectroscopic data created unified benchmarks that enabled systematic model comparison and community-wide progress [6, 13]. The scale of these datasets—often exceeding 100,000 entries—necessitated automated pipelines for data ingestion, cleaning, and featurization [4]. Transfer learning techniques further leveraged this abundance, allowing models pre-trained on abundant properties (e.g., formation energy) to be fine-tuned for scarcer targets (e.g., thermal conductivity) [4].

A key insight from integrative analysis is the synergistic role of computation and experiment. High-throughput virtual screening generates hypotheses at rates unattainable by manual effort, while targeted experiments validate and refine them [5, 12]. This feedback dynamic has been formalized in active learning frameworks, where model uncertainty guides the selection of the most informative new data points [18]. The resulting virtuous cycle has expanded known stable material spaces by orders of magnitude in recent years [12].

Representation learning: Capturing universal materials features

Traditional descriptor approaches, while effective, suffered from brittleness when applied outside their training distributions. Representation learning addressed this by shifting the burden of feature engineering to the model itself. Early neural network architectures learned embeddings from atomic coordinates and compositions, revealing latent spaces that clustered materials by functional similarity [6, 15].

Unsupervised methods, such as word-embedding analogs applied to chemical formulas or crystal structures, demonstrated that materials science literature itself encodes rich semantic relationships [15]. These embeddings captured concepts like “thermoelectric efficiency” or “superconducting transition temperature” without explicit supervision, providing a foundation for zero-shot property prediction.

Supervised representation learning further refined these embeddings by optimizing for multiple downstream tasks simultaneously [17]. Multi-task models not only improved individual property predictions but also exhibited emergent transferability: a model trained on elemental solids could generalize to complex oxides after minimal fine-tuning [4, 17]. This capability is a hallmark of foundation models and underscores the value of diverse, task-agnostic pre-training.

Graph neural networks: The architectural backbone

Graph neural networks have become the dominant paradigm for atomic-scale modeling in materials science [8–11]. Their message-passing architecture naturally respects the locality and hierarchy of chemical bonding, enabling accurate prediction of energies, forces, and electronic properties from first-principles reference data [9, 16].

Foundational works established that graph convolutions could achieve DFT-level accuracy for formation energies and band gaps on benchmark datasets [8, 9]. Subsequent architectures incorporated physical invariances—rotational equivariance, periodic boundary conditions, and long-range electrostatics—dramatically improving performance on dynamic and defective systems [10, 11].

Scaling laws observed in large GNNs reveal a critical principle: performance improves as a power law with both dataset size and model capacity [12]. When trained on millions of structures, these networks exhibit remarkable generalization, correctly identifying stable phases in chemical spaces never explicitly encountered during training [12]. This behavior suggests that sufficiently large and diverse graph-based pre-training can yield models that function as *de facto* foundation models for materials.

High-throughput computation and simulation–experiment integration

The computational backbone of data-driven materials engineering is high-throughput DFT and beyond-DFT methods. Automated workflows have generated petabyte-scale datasets of relaxed structures, phonon spectra, and defect energetics [5, 6, 13]. These data serve dual purposes: direct training of surrogate models and validation of experimental observations.

Integration of simulation and experiment occurs at multiple levels. Machine-learned interatomic potentials, trained on DFT forces, enable molecular dynamics simulations at experimental time- and length-scales [16]. Uncertainty quantification frameworks propagate errors from quantum calculations through surrogate models to experimental design [18]. The result is a tightly coupled ecosystem where computational predictions inform synthesis targets and experimental outcomes refine computational models [4, 21].

Active learning and uncertainty quantification

Active learning has matured into a cornerstone of efficient materials exploration. By quantifying epistemic uncertainty—arising from model limitations rather than data noise—researchers can prioritize experiments that maximally reduce predictive variance [18]. Bayesian neural networks and ensemble methods provide practical implementations, with acquisition functions balancing exploration (high-uncertainty regions) and exploitation (promising property spaces) [18, 24].

Cross-study synthesis reveals that uncertainty-aware pipelines consistently outperform random or greedy sampling, often achieving target properties with 5–10× fewer experiments [18]. When coupled with representation learning, these systems also identify previously unrecognized structure–property motifs, accelerating serendipitous discovery.

Inverse materials design

The ultimate goal of data-driven engineering is inverse design: specifying desired properties and obtaining candidate structures. Early generative models relied on variational autoencoders or generative adversarial networks operating in descriptor space [23]. Graph-based generative approaches have since enabled direct structure generation while enforcing physical constraints [23].

Foundation-model-inspired paradigms—pre-training on vast structural datasets followed by property-conditioned fine-tuning—promise to make inverse design routine. The combination of scalable GNN backbones, multimodal conditioning, and reinforcement learning from experimental feedback positions the field to solve long-standing inverse problems in catalysis, energy storage, and quantum materials.

Autonomous & closed-loop discovery systems

Autonomous laboratories represent the operational culmination of computational and data-driven advances. These systems integrate robotic hardware, orchestration software, and intelligent decision agents into closed-loop workflows that operate with minimal human intervention [19–22]. At their core are foundation-like models that synthesize knowledge from heterogeneous sources and direct experimental actions.

The architecture of a typical autonomous discovery platform comprises three interconnected layers. The perception layer aggregates data from in-situ characterization (X-ray diffraction, spectroscopy, microscopy) and high-throughput simulations [4, 21]. The reasoning layer—embodied by large-scale GNNs or multimodal transformers—generates hypotheses, ranks candidates, and quantifies uncertainty [12, 18]. The action layer executes synthesis, processing, and measurement protocols via robotic platforms [22].

Recent implementations illustrate the power of this integration. Self-driving platforms for thin-film optimization have discovered novel compositions with superior optoelectronic properties by iteratively refining deposition parameters based on real-time feedback [21]. Similarly, autonomous solid-state synthesis systems have accelerated the discovery of battery electrolytes by coupling robotic pipetting with machine-learned phase identification [22].

A defining feature of these systems is their use of active learning to navigate design spaces. Rather than exhaustively screening all possibilities, the model proposes experiments that maximize information gain, often formulated as expected improvement or upper confidence bound criteria [18, 24]. This approach has reduced the number of required syntheses from thousands to dozens for target optimization tasks [22].

Multimodal foundation models further enhance autonomy by fusing textual descriptions from literature, spectral signatures from experiments, and atomic configurations from simulations [15, 19]. Such models can, for instance, interpret a failed synthesis outcome in the context of prior DFT predictions and literature precedents, then propose corrective actions without human input. The operational integration of perception, reasoning, and action layers within autonomous discovery ecosystems is orchestrated through foundation-model cognition (Figure 2).

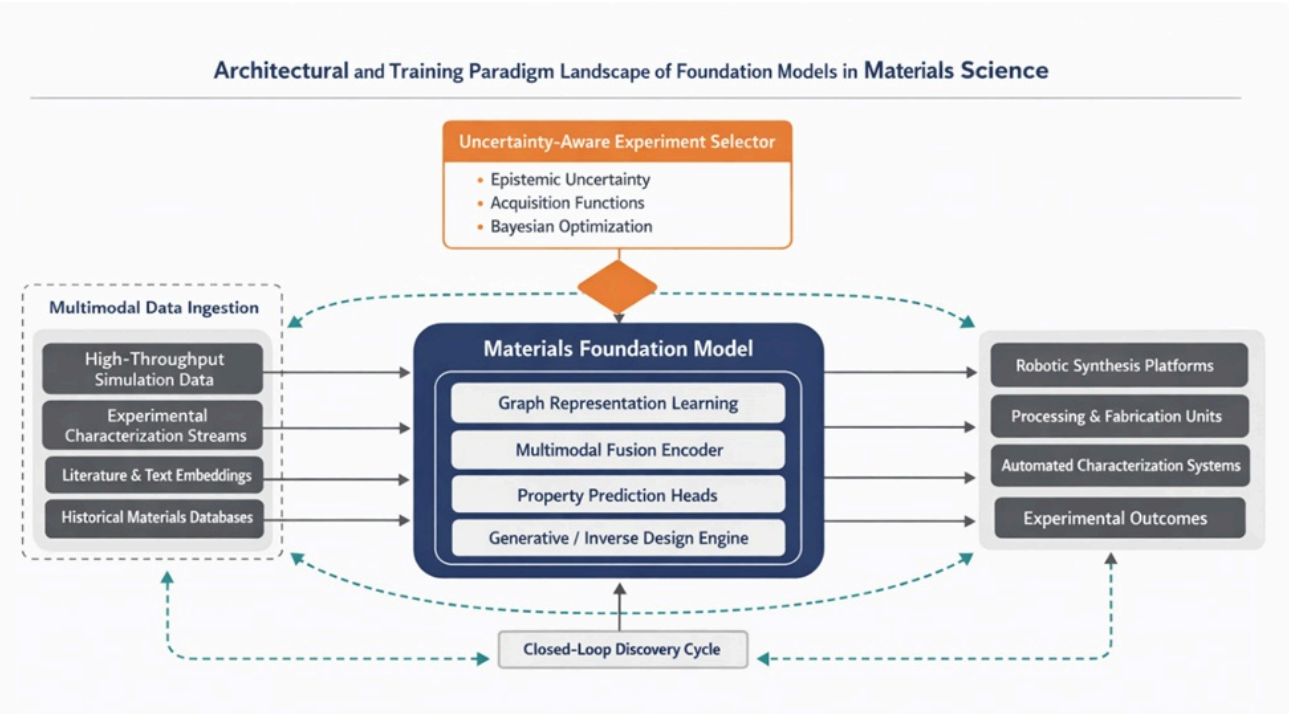


Figure 2. Foundation-Model-Driven Autonomous Closed-Loop Discovery Architecture

Systems architecture of an autonomous closed-loop materials discovery platform powered by foundation models. Multimodal perception streams—including high-throughput simulations, real-time experimental measurements, and literature-derived embeddings—are ingested into a large-scale foundation model that performs representation learning, property prediction, and inverse design. An uncertainty-aware active learning controller prioritizes experimental candidates and dispatches synthesis instructions to robotic execution systems. Experimental outcomes are reintegrated via continual learning loops, enabling adaptive model refinement and accelerated discovery.

The training paradigms that enable these systems emphasize continual and federated learning. Models are pre-trained on static datasets but fine-tuned in real time as new experimental data arrive [4]. This approach mitigates distribution shift and ensures that the model remains aligned with the evolving experimental reality. Uncertainty quantification plays a dual role: guiding experiment selection and providing confidence intervals for downstream decision-making [18].

Integrative analysis across recent demonstrations reveals common success factors. First, tight coupling between simulation and experiment prevents model drift. Second, hierarchical decision-making—strategic planning at the foundation model level and tactical execution at the robotic level—enables scalability. Third, open data standards and modular software architectures facilitate community contributions and rapid iteration [14, 19].

These systems are already delivering tangible impact. Autonomous platforms have identified new perovskite compositions for photovoltaics, discovered metastable phases for quantum computing, and optimized alloy compositions for additive manufacturing—all within weeks rather than years [21, 22]. As foundation models mature, their deployment in autonomous laboratories will further amplify these gains, enabling discovery at the scale and speed required to address global challenges in energy, sustainability, and health.

Results and Discussion

Emergent architectural principles for generalizable foundation models

The convergence of graph neural networks, representation learning, and large-scale pre-training has crystallized a set of architectural principles that distinguish emerging foundation models from earlier machine-learning approaches in materials science [8–12]. At the core is the explicit encoding of relational structure through message-passing mechanisms that propagate local chemical environments into global representations. This design choice has proven remarkably robust across length scales, from molecular fragments to defective crystals and interfaces [10, 11].

Recent scaling studies reveal that performance follows predictable power-law relationships with both dataset size and model capacity, a hallmark of foundation-model regimes [12]. When trained on hundreds of thousands to millions of DFT-relaxed structures, graph architectures begin to exhibit zero-shot generalization to entirely new chemical families, suggesting that sufficiently diverse pre-training can internalize the periodic table's underlying grammar [12, 14]. Equivariant and long-range-aware variants further enhance this capability by respecting physical symmetries that earlier descriptor-based models could only approximate [10, 11].

Cross-study synthesis identifies a second unifying principle: hierarchical feature extraction. Lower layers capture local bonding motifs, while deeper layers assemble these into mesoscale and functional descriptors—exactly the multi-scale reasoning required for realistic materials design [8, 9, 16]. This hierarchy mirrors the organization of materials knowledge itself and explains why foundation-like models succeed where task-specific networks plateau.

Training paradigms that balance scale, robustness, and physical grounding

Training strategies have undergone a parallel evolution toward paradigms that maximize data efficiency and generalization. Self-supervised objectives—particularly contrastive and masked reconstruction—allow models to learn from the intrinsic structure of unlabeled computational and experimental data [15, 16]. These approaches are especially potent in materials, where high-quality labels are expensive but raw structures and trajectories are abundant [15].

Multi-task learning provides an orthogonal axis of improvement by forcing shared representations to satisfy multiple physical constraints simultaneously [4, 17]. The resulting embeddings encode latent correlations (e.g., between formation energy, band gap, and elastic moduli) that improve downstream performance even on properties never explicitly trained [17]. When combined with active learning, these paradigms create self-reinforcing cycles in which uncertainty quantification directly guides data acquisition [18].

A third, more recent development is the integration of multimodal and continual learning signals [4, 15, 19]. Models that ingest atomic graphs, spectroscopic fingerprints, textual synthesis protocols, and literature embeddings in a unified latent space exhibit superior robustness to distribution shifts—the very shifts encountered when moving from simulation to

experiment [15, 19]. This multimodal fusion is not merely additive; it creates emergent capabilities such as literature-conditioned inverse design and failure-mode diagnosis that single-modality models cannot achieve [14, 15].

Autonomous systems as living laboratories for foundation-model validation

The true test of any foundation model is its performance inside a closed discovery loop. Autonomous laboratories have become the ultimate validation platforms because they enforce real-time feedback, physical constraints, and economic realities that static benchmarks cannot replicate [18–22]. In these systems, foundation models function simultaneously as predictor, planner, and critic: they propose candidates, select experiments via uncertainty-aware acquisition, interpret outcomes, and update their parameters on the fly [18, 21, 22].

Integrative analysis across recent deployments reveals three recurring success factors. First, tight simulation–experiment coupling prevents model drift and maintains physical consistency [4, 21]. Second, hierarchical control—strategic reasoning at the foundation-model level and tactical execution at the robotic level—enables scalability to hundreds of parallel experiments [19, 22]. Third, open data standards and modular orchestration frameworks accelerate community adoption and iterative improvement [14, 19].

These platforms are already demonstrating order-of-magnitude accelerations in materials optimization [21, 22]. More importantly, they are generating the high-fidelity, multimodal datasets that will fuel the next generation of even larger foundation models, creating a virtuous cycle of infrastructure and intelligence.

Challenges and limitations

Despite substantial progress, several fundamental limitations constrain the current generation of foundation models in materials science.

Data heterogeneity remains the most pervasive challenge. Computational datasets are orders of magnitude larger than experimental ones, yet the two domains differ systematically in noise characteristics, coverage, and metadata quality [4, 6, 7]. Transfer learning mitigates but does not eliminate this gap, and models trained predominantly on DFT data often exhibit optimistic bias when deployed in laboratory settings [4, 18].

Computational cost and energy consumption of large-scale training and inference constitute a second bottleneck. Training a state-of-the-art graph foundation model can require thousands of GPU-hours, raising questions of accessibility and environmental sustainability [12, 14]. Inference latency, while dramatically lower than DFT, can still be prohibitive for real-time robotic control in high-throughput platforms [19, 22].

Interpretability and physical consistency present a deeper scientific limitation. Even the most accurate models remain largely black-box, complicating the extraction of mechanistic insight and the detection of physically implausible predictions [10, 16, 18]. Physics-informed constraints help, but enforcing them at scale without sacrificing expressivity remains an open architectural problem [11, 16].

Generalization to complex, non-equilibrium, and multi-component systems is still limited. Most current foundation models excel on crystalline solids and small molecules but struggle with amorphous materials, interfaces, defects at finite temperature, and multi-phase microstructures—precisely the regimes of greatest technological relevance [10–12].

Finally, the field lacks standardized benchmarks and evaluation protocols tailored to foundation-model capabilities. Existing leaderboards emphasize narrow property prediction tasks and do not capture the broader requirements of inverse design, uncertainty calibration, or continual adaptation that define real-world deployment [6, 14].

Future research directions

The next decade of research will likely be defined by three strategic thrusts that directly address the limitations above while extending the foundation-model paradigm.

First, the construction of truly multimodal, continuously updated foundation models that ingest and align data from literature, simulation, spectroscopy, microscopy, and robotic experimentation in a single shared representation space [4, 19, 25]. Federated and privacy-preserving training protocols will be essential to aggregate experimental data across institutions without compromising intellectual property [26].

Second, the development of hybrid architectures that deeply embed physical laws—through equivariant operators, differentiable simulators, and learned interatomic potentials—into the very fabric of foundation models [7, 19, 27]. These physics-augmented models will not only improve accuracy but also enable reliable extrapolation far beyond training distributions and provide interpretable mechanistic insights [16, 17].

Third, the maturation of autonomous discovery ecosystems into self-improving scientific agents. This will require advances in meta-learning (models that learn how to learn), reinforcement learning over long horizons, and human–AI collaborative interfaces that preserve scientific agency while leveraging machine scale [18–20, 22]. Particular emphasis should be placed on uncertainty-aware planning that balances exploration of novel chemical space with exploitation of known high-value regions [18, 24].

Additional promising directions include the creation of community-wide foundation-model benchmarks that evaluate generalization, robustness, and discovery efficiency; the integration of causal inference techniques to move beyond correlation toward mechanistic understanding; and the systematic study of scaling laws specific to materials science to guide efficient resource allocation [12, 14].

Conclusion

Foundation models are emerging as the computational substrate for a new era of materials science—one in which the design, prediction, synthesis, and optimization of materials occur at unprecedented speed and scale. The architectures, training paradigms, and autonomous systems reviewed here collectively form a coherent infrastructure that is already compressing decades-long discovery timelines into months.

By synthesizing advances across materials informatics, graph representation learning, high-throughput computation, and closed-loop experimentation, this review has highlighted the integrative principles that will define the next generation of foundation models: hierarchical relational reasoning, multimodal fusion, uncertainty-aware decision making, and continual adaptation grounded in physical reality.

The path forward is clear. If the materials community invests in the shared data ecosystems, open architectures, and collaborative platforms required to train and deploy these models at scale, the coming decade will witness a materials renaissance capable of addressing grand challenges in energy, sustainability, quantum technologies, and human health. The foundation has been laid; the task now is to build upon it.

Acknowledgements

None

Conflict of interest

None

Financial support

None

Ethics statement

None

Received: 02 Apr 2024

Revised: 05 May 2024

Accepted: 04 Jun 2024

Published online: 18 September 2024

Rights and permissions

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Ramprasad R, Batra R, Pilania G, Mannodi-Kanakkithodi A, Kim C. Machine learning in materials informatics: Recent applications and prospects. *npj Comput Mater.* 2017;3:54.
- Butler KT, Davies DW, Cartwright H, Isayev O, Walsh A. Machine learning for molecular and materials science. *Nature.* 2018;559(7715):547-55.
- Choudhary K, DeCost B, Tavazza F. Machine learning with force-field inspired descriptors for materials: Fast screening and mapping energy landscapes. *Phys Rev Mater.* 2018;2(8):083801.
- Jha D, Choudhary K, Tavazza F, Liao WK, Choudhary A, Campbell C, et al. Enhancing materials property prediction by leveraging computational and experimental data using deep transfer learning. *Nat Commun.* 2019;10(1):5316.
- Pilania G, Gubernatis JE, Lookman T. Multi-fidelity machine learning models for accurate bandgap predictions of solids. *Comput Mater Sci.* 2017;129:156-63.
- Himanen L, Geurts A, Foster AS, Rinke P. Data-driven materials science: Status, challenges and perspectives. *Adv Sci.* 2019;6(21):1900808.
- Schmidt J, Marques MRG, Botti S, Marques MAL. Recent advances and applications of machine learning in solid-state materials science. *npj Comput Mater.* 2019;5:83.
- Chen C, Ye W, Zuo Y, Zheng C, Ong SP. Graph networks as a universal machine learning framework for molecules and crystals. *Chem Mater.* 2019;31(9):3564-72.
- Xie T, Grossman JC. Crystal graph convolutional neural networks for an accurate and interpretable prediction of material properties. *Phys Rev Lett.* 2018;120(14):145301.

Reiser P, Neubert M, Eberhard A, Torresi L, Zhou C, Shao C, et al. Graph neural networks for materials science and chemistry. *Commun Mater.* 2022;3(1):93.

Fung V, Zhang J, Juarez E, Sumpter BG. Benchmarking graph neural networks for materials chemistry. *npj Comput Mater.* 2021;7:84.

Merchant A, Batzner S, Schoenholz SS, Aykol M, Cheon G, Cubuk ED. Scaling deep learning for materials discovery. *Nature.* 2023;624(7990):80-5.

Li Y, Mao W, Wang G, Liu J, Wang S. A general-purpose machine learning framework for predicting singular integrals in boundary element method. *Eng Anal Bound Elem.* 2020;117:41-56.

Batra R, Song L, Ramprasad R. Emerging materials intelligence ecosystems propelled by machine learning. *Nat Rev Mater.* 2021;6(8):655-78.

Tshitoyan V, Dagdelen J, Weston L, Dunn A, Rong Z, Kononova O, et al. Unsupervised word embeddings capture latent knowledge from materials science literature. *Nature.* 2019;571(7763):95-8.

Xie T, France-Lanord A, Wang Y, Shao-Horn Y, Grossman JC. Graph dynamical networks for unsupervised learning of atomic scale dynamics in materials. *Nat Commun.* 2019;10(1):2667.

Janet JP, Kulik HJ. Predicting electronic structure properties of transition metal complexes with neural networks. *Chem Sci.* 2017;8(7):5137-52.

Lookman T, Balachandran PV, Xue D, Yuan R. Active learning in materials science with emphasis on adaptive sampling using uncertainties for targeted design. *npj Comput Mater.* 2019;5:21.

Tom G, Schmid SP, Baird SG, Cao Y, Darvish K, Hao H, et al. Self-driving laboratories for chemistry and materials science. *Chem Rev.* 2024;124(16):9633-732.

Häse F, Roch LM, Aspuru-Guzik A. Next-generation experimentation with self-driving laboratories. *Trends Chem.* 2019;1(3):282-91.

MacLeod BP, Parlange FGL, Morrissey TD, Häse F, Roch LM, Dettelbach KE, et al. Self-driving laboratory for accelerated discovery of thin-film materials. *Sci Adv.* 2020;6(20):eaaz8867.

Szymanski NJ, Rendy B, Fei Y, Kumar A, Banik A, Wang Y, et al. An autonomous laboratory for the accelerated synthesis of novel materials. *Nature.* 2023;624(7990):86-91.

Zunger A. Inverse design in search of materials with target functionalities. *Nat Rev Mater.* 2018;3(4):1-16.

Häse F, Roch LM, Kreisbeck C, Aspuru-Guzik A. Phoenix: A Bayesian optimizer for chemistry. *ACS Cent Sci.* 2018;4(9):1134-45.

Jha D, Gupta V, Ward L, Yang Z, Wolverton C, Foster I, et al. Enabling deeper learning on big data for materials informatics applications. *Sci Rep.* 2021;11(1):4244.

Morgan D, Jacobs R. Opportunities and challenges for machine learning in materials science. *Annu Rev Mater Res.* 2020;50:71-103.

Kauwe SK, Graser J, Murdock R, Sparks TD. Can machine learning find extraordinary materials? *Comput Mater Sci.* 2020;174:109499.